International Journal of Mathematics and
Computer Applications Research (IJMCAR)
ISSN(P 2249-6955; ISSN(E): 2249-8060
Vol. 5, Issue 4, Aug 2015, 83-92
TJPRC Pvt. Ltd.

CONDITION FOR FESIABLE SOLUTION TO PRIMAL AND DUAL NONLINEAR PROGRAMMING PROBLEM

PRASHANT CHAUHAN

S. L. Education Institute, Moradabad, Uttar Pradesh, India

ABSTRACT

In this paper, we take a quick look at some results that have been worked out for a kind of duality theory for nonlinear programming problems. These are quite parallel to those in the linear programming case.

KEYWORDS: Nonlinear Programming, Duality, Lagrangian Function

INTRODUCTION

The general nonlinear programming maximization problem is

Maximize $\quad II(X) = f(X)$

Subject to $\quad h(x) \leq 0$ (1)

And $\quad X \geq 0$.

Where, as usual $h(x) = [h^1(X), \dots, h^m(X)]'$. the associated Lagrangian function, used primarily in development of Kuhn-Tucker [18] conditions and in the discussion of saddle points, is

$$L(X,Y) = f(X) - Y'[h(X)] \tag{2}$$

Where $Y = \left[y_1, \dots, y_m\right]'$. The gradient of L(X,Y) with respect to the y's is just $\nabla L_Y = -h(X)$, and so f(X) in (1) could be expressed as

$$f(X) = L(X,Y) + Y'[h(X)] = L(X,Y) - Y'\nabla L_Y$$

In addition, the requirements in $h(X) \leq 0$ can be expressed as $\nabla L_Y \geq 0$, so the general nonlinear programming maximization problem in (1) could be also be written

Maximize $\; L(X,Y) - Y'\nabla L_Y$

Subject to $\; -\nabla L_Y \leq 0$ (3)

And $\quad X \geq 0$

This appears only to add complexity to the expression in (1), but it does suggest a "symmetric" problem:

Minimize $L(X,Y) - X'\nabla L_X$

Subject to $-\nabla L_X \geq 0$ (4)

And $Y \geq 0$

From L(X, Y) in (2), and using the Jacobian matrix

$$J = \begin{bmatrix} [\nabla h^1]' \\ . \\ . \\ . \\ . \\ . \\ . \\ [\nabla h^m]' \end{bmatrix}$$

We have

$$\nabla L_X = \nabla f - J'Y$$

And so (4) in more detail is

Minimize $\Delta(X,Y) = f(X) - Y'[h(X)] - X'[\nabla f - J'Y]$

Subject to $J'Y - \nabla f \geq 0$ (5)

And $Y \geq 0$

Written out in more detail this is

Minimize $f(X) - \sum_{i=1}^{m} y_i h^i(X) - \sum_{j=1}^{n} x_j (f_j - \sum_{i=1}^{m} y_i h_j^i)$

Subject to $f(X) - \sum_{i=1}^{m} y_i h_j^i - f_j \geq 0 \qquad (j = 1, \ldots.., n)$ (6)

And $y_i \geq 0 \qquad (i = 1, \ldots.., m)$

When f(X) in the maximization problem is concave and each constraint is convex or quasiconvex, so that both necessary and sufficient conditions for a maximum to the nonlinear programming problem, the (6) [or(5) is taken to be the dual to (1) or (3)]. This is primarily because (I) a pair of dual linear programs corresponds precisely to (3) and (4), and (II) a set of theorems parallel to those in duality theory for linear programs can be derived.

Regarding the linear programming connection, the general linear programming maximization problem was

Maximize $P'X$

Subject to $AX \leq B$

And $X \geq 0$

The associated Lagrangian function, as in (2) would be

$$L(X,Y) = P'X - Y'(AX - B) \tag{7}$$

$$\nabla L_X = [P - A'Y]$$

$$P'X - Y'[AX - B] - X'[P - A'Y]$$

$$AX \leq B0X \geq 0$$

And so $\nabla L_Y = -[AX - B]$ and $\nabla L_X = [P - A'Y]$. (The latter gradient is defined as a column vector to conform to the convention of expressing gradient as columns.) The primal linear program can now be written in the style of (3) as

Maximize $P'X - Y'[AX - B] + Y'[AX - B]$

Subject to $AX - B \leq 0$ or $AX \leq B$ (8)

And $X \geq 0$

Corresponding to (4) we have, for this linear program,

Minimize $P'X - Y'[AX - B] - X'[P - A'Y]$

Or, since $P'X = X'P, Y'B = B'Y$ and $Y'AX = X'A'Y$

Minimize $B'Y$

Subject to $-P + A'Y \geq 0$ or $A'Y \geq P$

And $Y \geq 0$.

Clearly, (7) and (8) are exactly a pair of dual linear programs.

We now explore a set of primal-dual theorems for the nonlinear programming problems in (1) and (5). We use a "prime" for these in the nonlinear case; there are obvious parallels to the results of linear programs.

CONDITION FOR THE PRIMALAND DUAL PROBLEM TO HAVE FESIABLE SOLUTION

Theorem 1

Feasible solutions to the primal and dual problems are optimal if and only if, $II(X^p) = \Delta(X^d Y^d)$.

Clearly, if $II = \Delta$, then both objective functions have reached their limits and so the solutions are optimal. An

important outcome of the proof is t6haty, given an optimal X* for the primal, a vector Y* can be found such that (X*, Y*) is an optimal solution to the dual.

Theorem 2

A pair of feasible solutions has $II(X^*) = \Delta(X^*, Y^*)$ if and only if (I) $(Y^*)'[-h(X^*)] = 0$ and (II) $(X^*)'[\nabla f^* - (J^*)'Y^*] = 0.$

Note that $[-h(X^*)]$ in (I) is the vector of slack variables for the constraints in (1), and $[\nabla f^* - (J^*)'Y^*]$ in (5). So this theorem describes a kind of “complementary slackness” for optimal solutions that is parallel to that in the linear programming case. This property of the optimal solutions to the pair of dual nonlinear programs shows the equivalence of the dual variables (Y) and Lagrange multipliers the development of Kuhn-Tucker conditions and the saddle point problem connection.

Theorem 3

Under certain conditions, $\partial II(X^*)/\partial b_i = y_i^*$. This marginal valuation property of the dual variables at optimum held for linear programs, subject to the qualification that the appropriate derivatives existed. The same sort of requirement is needed here. Under those conditions y_i^*, is a measure of the impact on the4 optimum value of the primal objective function of a marginal change in b_i.

The Solution of Nonlinear Equations

There are many numerical methods which exist for locating the roots. We present on simple method here. The method presented here is called Newton’s method and is motivated as follows.

We assume that f has continuous second derivatives and that some estimate x_1 of a solution to

$$f'(x) = 0 \tag{9}$$

Is a available. If no such estimate is known, x_1 is chosen at random. If x_1 is a reasonably good estimate, the Taylor series expansion of f' about x_1 can be approximated as:

$$f'(x) = f'(x_1) + (x - x_1) f''(x_1).$$

Hence if x is a solution to (9),

$$0 = f'(x_1) + (x - x_1) f''(x_1) .$$

$$x = x_1 - f'(x_1) / f''(x_1) \tag{10}$$

Now unless f is a quadratic, x will not in general be an exact solution to (9). However, x can be used as an improved estimate. Indeed, (10) can be looked upon as the first equation in a family which generates successive improved estimates of a solution to (9). The family has the following general form:

$$x_{n+1} = \left[x_n - \frac{f'(x_n)}{f''(x_n)} \right]. \qquad \text{n} = 1,2,... \tag{11}$$

Once an estimate is finally found which is sufficiently close to a root, a new starting point can be selected in an effort to find a new root. This procedure is repeated until all roots are found.

Multidimensional Optimization with Equality Constraints

This problem can be stated in general as follows

$$\text{Maximize} \quad f(X) \tag{12}$$

$$\text{subject to} \qquad \mathrm{g_j(X) = 0.} \qquad \mathrm{j = 1,2,...., m} \tag{13}$$

Where

$$\mathrm{X = (x_1, x_2,....x_n)^T}$$

Consider the problem:

$$\text{Maximize} \qquad f(X) = x_1^2 + 2(x_2 - 4)^2 + 8$$

$$x_1 = x_2^2 - 4.$$

We are left with following unconstrained problem in one dimension:

$$\text{Maximize} \qquad f(x_1) = (x_2^2 - 4)^2 + 2(x_2 - 4)^2 + 8,$$

Which is easier to solve. Of course, this approach of elimination will be successful in reducing the number of variables in the problem only if it is possible to express a solution for one or more of the variables explicitly. Often, however, this cannot be done.

The Jacobian Method

We now present a method which solves the problem (12), (13).It is assumed that f and g_j, $j = 1,2,3,......,m.$ has continuous second derivatives. The strategy is to find a suitable expression for the first derivatives of f at all points which satisfy (13). The feasible stationary points of f are the ones among these for which

$$\frac{\partial f}{\partial x_i} = 0 \qquad \mathrm{i = 1,2,..., n} \tag{14}$$

The maximum points are identified among those satisfying (14).

These ideas are now placed on a firm mathematical basis. Consider any point *X* which satisfies (13). In any neighborhood of *X* there will exist at least one point *X+h* which satisfies (13), because *X* is on the boundary of the region defined by (13). Expanding *f* and g_j, $j = 1,2,...,m$, in a Taylor series about *X*, we get

$$f(X+h) = f(X) + \nabla f(X)^T h + \frac{1}{2} h H_f(\theta X + (1-\theta)(X+h))h,$$

$$g_j(X+h) = g_j(X) + \nabla g_j(X)^T h + \frac{1}{2} h H_{g_j}(\theta X + (1-\theta)(X+h))h, \qquad j = 1,2,\dots,m,$$

For some $\theta, 0 < \theta < 1$ .As $X+h$ approaches X, we get

$$f(X+h) \approx f(X) + \nabla f(X)^T h$$
$$g_j(X+h) \approx g_j(X) + \nabla g_j(X)^T h \qquad j = 1,2,..,m$$

Therefore

$$\partial f(X) \approx \nabla f(X)^T \partial X$$
$$\partial g_j(X) \approx \nabla g_j(X)^T \partial X \qquad j = 1,2,\dots,m,$$

Using (13) we get

$$\partial g_j(X) = 0 \qquad j = 1,2,\dots,m.$$

Thus we can state, to within a first order approximation

$$\nabla g_j(X)^T \partial X = 0, \qquad \mathrm{j} = 1,2,\dots,\mathrm{m} \tag{15}$$

Now as $\nabla f(X)$ and $\nabla g_j(X),\ \mathrm{j} = 1,2,\dots,\mathrm{m}$ consist of known constants, (15) constitutes a set of (m+1) linear equations in (n+1) unknowns, $\partial x_1, \partial x_2, \partial x_3, \dots, \partial x_m, \partial f(X)$ .If the equations are linearly dependent one discards the smallest number whose removal leaves an independent set. Hence we can assume that there are no more equations than variables, i.e.

$$m \le n$$

Now

$m=n$

Leads to the unique solution

$$\partial X = 0$$

Which implies that there are no feasible points other than X in any neighbourhood of X. That is, the set of feasible points is discrete. Hence, we can assume that

$m<n$.

We redefine $X = (x_1, x_2, \dots\dots x_n)^T$ as

$$X = (w_1, w_2, ..., w_m, y_1, y_2, ...y_{n-m})^T \tag{16}$$

The variables $w_i, i = 1,2,...,m$ are called state variables and the variables $y_i, i = 1,2,...,(n-m)$ are called decision variables. Now (15) can be rewritten using (16), as follows:

$$\sum_{i=1}^{m} \frac{\partial f(X)}{\partial w_i} \partial w_i + \sum_{i=1}^{n-m} \frac{\partial f(X)}{\partial y_i} \partial y_i = \nabla f(X) \tag{17}$$

$$\sum_{i=1}^{m} \frac{\partial g_j(X)}{\partial w_i} \partial w_i + \sum_{i=1}^{n-m} \frac{\partial g_j(X)}{\partial y_i} \partial y_i = 0 \qquad j = 1,2,...,m \tag{18}$$

Suppose now that the $\partial y_i, i = 1,2,..., (n-m)$ are given arbitrary values. When these are substituted into unique values for the $\partial w_i, i = 1,2,...,m$ can be found which keep X+h inside the feasible region. One can then use all these values in (17) to see if

$$\partial f(X) > 0$$

i.e., the new point X+h is an improvement over X.

We now state the explicit steps needed to carry this out using vector notation. The matrix

$$\begin{pmatrix} \frac{\partial g_1}{\partial w_1} & \frac{\partial g_1}{\partial w_2} & \cdots & \frac{\partial g_1}{\partial w_m} \\ \frac{\partial g_2}{\partial w_1} & \frac{\partial g_2}{\partial w_2} & \cdots & \frac{\partial g_2}{\partial w_m} \\ \vdots & \vdots & & \vdots \\ \frac{\partial g_m}{\partial w_1} & \frac{\partial g_m}{\partial w_2} & \cdots & \frac{\partial g_m}{\partial w_m} \end{pmatrix}$$

Is called the Jacobian matrix, and the matrix

$$C = \begin{pmatrix} \frac{\partial g_1}{\partial y_1} & \frac{\partial g_1}{\partial y_2} & \cdots & \frac{\partial g_1}{\partial y_{n-m}} \\ \frac{\partial g_2}{\partial y_1} & \frac{\partial g_2}{\partial y_2} & \cdots & \frac{\partial g_2}{\partial y_{n-m}} \\ \vdots & \vdots & & \vdots \\ \frac{\partial g_m}{\partial y_1} & \frac{\partial g_m}{\partial y_2} & \cdots & \frac{\partial g_m}{\partial y_{n-m}} \end{pmatrix}$$

Is called the control matrix. It is important in defining the state and decision variables that the left-hand sums in (17) and (18) be linearly independent. It is always possible to make a choice of which x_i's become state variables. So this happens because we have assumed that the equations in (15) are linearly independent. The implication of this is that J is nonsingular. Now let

$$W = (w_1, w_2, ..., w_m)^T$$

$Y = (y_1, y_2, ...y_{n-m})^T$

Then

$$\nabla_w f^T \partial W + \nabla_w f^T \partial Y = \partial f(W, Y) \tag{19}$$

And

$$J\partial W + C\partial Y = 0 \tag{20}$$

Respectively. As J is nonsingular, we can multiply (20) by J^{-1}.

$$\partial W = -J^{-1} C \partial Y \tag{21}$$

It can be seen, that if the elements in ∂Y are given values, ∂W can be calculated using (21). Substituting this into (19) yields

$$\partial f(W, Y) = (\nabla_y f^T - \nabla_w f^T J^{-1} C)\partial Y \tag{22}$$

From (22) we can form what is known as the constrained gradient of f with respect to Y, which is

$$\nabla_y^c f = \frac{\partial^c f(w, y)}{\partial^c y} = \nabla_y f^T - \nabla_w f^T J^{-1} C \tag{23}$$

Each element of $\nabla_y^c f$, namely $\dfrac{\partial^c f}{\partial^c y_i}, i = 1,2,...(n-m)$ is called a constrained derivative. It represents the rate of change of f resulting from perturbing x_i from y_i (all other x_i 's being held constant) to feasible points.

When constrained derivatives are used, i.e. X^* is a feasible maximum it is necessary that

$$\nabla_y^c f(X^*) = 0 \tag{24}$$

Equation (24) can be used identify all the stationary points; it remains to find which one is the global maximum. With the modification that H is the matrix of constrained second derivatives with respect to the independent variables $y_1, y_2, ..., y_{n-m}$ only, and not $w_1, w_2, ..., w_m$. The complete method will be illustrated with a numerical example.

CONCLUSIONS

There are two principal reasons for interest in nonlinear duality at this point:(1) for any feasible solutions to a pair of primal and dual nonlinear programming problems, the dual objective function value provides a limit on the value of the primal objective function (as with linear programs) and (2) for a pair of optimal solutions, the value of the dual variables may have the same kind of “shadow price” interpretation that we associate with the linear programming case-giving a possible marginal valuation to resource that are used up in the optimal solution and a value of zero to those resources that are in excess supply at an optimal solution.

REFERENCES

1. **Andreani, R., and Marti Nez, J.M.,** On the Solution of Mathematical Programming Problems with Equilibrium Constraints, Mathematical Methods of Operations Research, Vol. 54, pp. 345–358, 2001.
2. **Bertsekas, D.P.,** Nonlinear Programming, 2nd Edition, Athena Scientific,Belmont, Massachusetts, 1999.
3. **Bigi, G., and Pappalardo, M**., Regularity Conditions in Vector Optimization, Journal of Optimization Theory and Applications, Vol. 102, pp. 83–96, 1999.
4. **Boggs, P. T., and Tolle, J.W**., A Strategy for Global Convergence in a Sequential Quadratic Programming Algorithm, SIAM Journal on Numerical Analysis, Vol. 26, pp. 600–623, 1989.
5. **Castellani,M.,** Separation of Sets and Optimality Conditions, Rendiconti del Circolo Matematico di-Palermo, Series II, Vol. 48, pp. 27–38, 1997.
6. **Dennis, J.E., El-Alem, M. and Maciel, M.C.,**A Global Convergence Theory for General Trust-Region-Based Algorithms for Equality Constrained Optimiz-ation,SIAM Journal on Optimization, Vol. 5, pp. 177–207, 1995.
7. **Dien,P.H.,Mastroeni,G,Pappalardo,M.,andQUANG,P.H.,** Regularity Conditions for Constrained Extremum Problems in Image Space: The Nonlinear Case, Journal of Optimization Theory and Applications, Vol. 80, pp. 19–37,1994.
8. **Facchinei, F., Jiang, H., and QI, L.,** A Smoothing Method for Mathematical Programs with Equilibrium Constraints, Mathematical Programming, Vol. 85,pp. 107–134, 1999.
9. **Favati, P., and Pappalardo, M.,** On the Reciprocal Vector Optimization Problem, Journal of Optimization Theory and Applications, Vol. 47, pp. 181–193, 1985.
10. **Fernandes, L., Friedlander, A.,Guedes, M., and JU ´ Dice, J. J.,** Solution of Generalized Linear Complementarity Problems Using Smooth Optimization and Its Application to Bilinear Programming and LCP, Applied Mathematics and Optimization, Vol. 43, pp. 1–19, 2001.
11. **Fiacco, A. V., and Mccormick, G. P.,** Nonlinear Programming, John Wiley,New York, NY, 1968.
12. **Giannessi, F.,** General Optimality Conditions in a Separation Scheme, Algorithms for Continuous Optimization, Edited by E. Spedicato, Kluwer, Dord-recht, Netherlands, pp. 1–23, 1994.
13. **Giannessi, F., and Rapcsak, T.,** Images, Separation of Sets, and ExtremumProblems, Recent Trends in Optimization Theory and Applications, Edited by R. P. Agarwal, Series in Applicable Analysis, World Scientific, Singapore,pp. 79–106, 1996.
14. **Giannessi,F.,** On Connections among Separation, Penalization, and Regularization for variational inequalities with Point-to-Set Operators, Rendiconti del Cir-colo Matematico di Palermo, Series` II, Vol. 48, pp. 137–145, 1997.
15. **Giannessi, F., Mastroeni, G., and Pellegrini, L.**, On the Theory of Vector Optimization and Variational Inequalities: Image Space Analysis and Separation, Vector Variational Inequalities and Vector Equilibrium, Mathematical Theories, Edited by F. Giannessi, Kluwer, Dordrecht, Netherlands, pp. 153–216, 2000.

16. **Giannessi, F., and Pellegrini, L.,** Image Space Analysis for Vector Optimization and Variational Inequalities: Scalarization, Advances in Combinatorial and Global Optimization, Edited by A. Migdalas, P. Pardalos, and R. Burkard, World Scientific, River Edge, pp. 982–1121, 1999.

17. **John, F.,** Extremum problems with inequalities as subsidiary conditions, Studies and Essays, courant anniversary volume Wiley-Interscience, New York, 187-204, 1948.

18. **Kuhn H.W. and Tucker A.W.**, "Non-linear programming",proceeding of second Berkley symposium of mathematical statistics and probability,481-492,1951.

19. **Mangasarian, O. L., and Fromovitz, S.,** The Fritz–John Necessary Optimality Conditions in Presence of Equality and Inequality Constraints, Journal of Mathematical Analysis and Applications, Vol. 17, pp. 37–47, 1967.

20. **Marti, Nez, J. M.,** Two-Phase Model Algorithm with Global Con convergence for Nonlinear Programming, Journal of Optimization Theory and Applications, Vol. 96, pp. 397–436, 1998.

21. **Marti, Nez, J. M., and Pilotta, E.A.,** Inexact Restoration Algorithms for Con-strained Optimization, Journal of Optimization Theory and Applications, Vol. 104, pp. 135–163, 2000.

22. **Mond B. and Weir T..,** Generalized concavity and duality," in Generalized concavity in optimization and economics (eds. S. Schaible and W.T. Ziemba), (Academic Press 1981), 263-279.

23. **Pappalardo,M.,**On Stationarity in Vector Optimization Problems, Rendicontidel Circolo Matematico di Palermo, Serie II, Supplemento 48, pp. 195–201,1997.

24. **Pappalardo, M.,** Vector Optimization, Encyclopedia of Optimization, Edited by P. Pardalos and C. Floudas, Kluwer, Dordrecht, Netherlands (to appear).

25. **Psiaki, M. L., and Park, S.,** Augmented Lagrangian Nonlinear Programming Algorithm that Uses SQP and Trust Region Techniques, Journal of Optimization Theory and Applications, Vol. 86, pp. 311–326, 1995.

26. **Rockafellar, R.T.,** Lagrange Multipliers and Optimality, SIAM Review, Vol.35, pp. 183–238, 1993.

27. **Spellucci, P.,** An SQP Method for General Nonlinear Programs Using Only Equality Constrained Sub problems, Mathematical Programming, Vol. 82, pp. 413–447, 1998.

28. **Vicente, L. N., and Calamai, P.,** A Bibliography, Journal of Global Optimization, Vol. 5, pp. 291–306, 1994.

29. **Zoutendijk, G.,** Methods of feasible directions, American Elsevier, New York, 1960.

30. **Zoutendijk, G.,** "Nonlinear programming, A numerical survey", SIAM J. control, 4, no. 1, Feb 1966.

31. **Zoutendijk, G.,** "Nonlinear programming,Computational methods," in Abadie[2],1970.